

\*\*\*\*\*

MPSRCH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Oct 6 04:39:24 1998; MasPar time 4667.32 Seconds  
1514.879 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-1  
Description: (3700-7548) from US08908884.seq (3 of 3)  
Perfect Score: 3849  
N.A. Sequence: 3700 GATGATATTGAGTTAGTCAA.....GGTGTGGCTAGTAGAAGCTT 7548  
Comp: CTACTATAACTCAATCAGTT.....CCACACCGATCATCTTCGAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532259 seqs, 918475165 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embl55  
1:em\_ba 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro

Database: genbank107  
12:gb\_ba 13:gb\_htg 14:gb\_in 15:gb\_om 16:gb\_ov 17:gb\_pat  
18:gb\_ph 19:gb\_pl 20:gb\_pr1 21:gb\_pr2 22:gb\_ro 23:gb\_st  
24:gb\_sts 25:gb\_sy 26:gb\_un 27:gb\_vi

Statistics: Mean 12.169; Variance 5.411; scale 2.249

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description	Pred. No.	
No.								
	1	1949	50.6	5655	19	ATU87794	Arabidopsis thaliana t	0.00e+00
	2	490	12.7	2104	19	ATU76707	Arabidopsis thaliana r	0.00e+00
c	3	101	2.6	79837	19	AB005233	Arabidopsis thaliana g	2.69e-58
	4	92	2.4	7218	17	I66494	Sequence 14 from paten	1.44e-50
c	5	52	1.4	7218	17	I66494	Sequence 14 from paten	5.55e-18
	6	39	1.0	10772	14	AF012089	Drosophila melanogaste	1.55e-08
c	7	39	1.0	10772	14	AF012089	Drosophila melanogaste	1.55e-08
	8	31	0.8	215	17	I28278	Sequence 5 from patent	2.65e-03
c	9	32	0.8	215	17	I28278	Sequence 5 from patent	6.36e-04
	10	31	0.8	354	14	OFU89259	Oxytricha fallax 57kD	2.65e-03
c	11	28	0.7	354	14	OFU89259	Oxytricha fallax 57kD	1.64e-01
c	12	26	0.7	370	14	OFU89262	Oxytricha fallax 57kD	2.20e+00
c	13	26	0.7	448	24	G29607	human STS SHGC-33999,	2.20e+00
c	14	26	0.7	37708	21	AC004034	Homo sapiens chromosom	2.20e+00
c	15	26	0.7	48889	14	CEY57G11B	Caenorhabditis elegans	2.20e+00
c	16	26	0.7	91017	21	HS162C6	Human DNA sequence fro	2.20e+00
c	17	26	0.7	110384	21	HSAC000117	Human BAC clone RG062A	2.20e+00
c	18	26	0.7	135672	21	HSAC000113	Human PAC clone DJ302C	2.20e+00
	19	27	0.7	194985	22	AC002406	Mouse chromosome X BAC	6.12e-01
c	20	25	0.6	60	17	A34729	(modified) eglin N-ter	7.65e+00
	21	25	0.6	323	24	G10153	human STS CHLC.GATA101	7.65e+00
c	22	24	0.6	354	14	AF036459	Trypanosoma cruzi muc	2.56e+01
c	23	24	0.6	375	19	DEZ99136	Drimia elata chloropla	2.56e+01
	24	25	0.6	565	17	E04076	gDNA encoding envelope	7.65e+00
c	25	25	0.6	882	27	HIVT1VI35	Human immunodeficiency	7.65e+00
c	26	25	0.6	1659	17	E08319	cDNA encoding lucifera	7.65e+00
	27	24	0.6	1663	15	MVU92534	Mustela vison microsat	2.56e+01
c	28	24	0.6	3853	16	CCNAKATP	C.commersoni mRNA for	2.56e+01
c	29	25	0.6	3989	27	EPVSPHER	Entomopoxvirus gene fo	7.65e+00
	30	24	0.6	5656	21	HSU49957	Human LIM protein (LPP	2.56e+01
c	31	24	0.6	6487	21	HSA2MGLB1	H.sapiens gene for alp	2.56e+01
c	32	24	0.6	30758	13	CEH06I04	Caenorhabditis elegans	2.56e+01
	33	25	0.6	33274	14	CEF52B5	Caenorhabditis elegans	7.65e+00
c	34	24	0.6	37738	14	CELF32B5	Caenorhabditis elegans	2.56e+01
	35	25	0.6	43589	14	CELC54D2	Caenorhabditis elegans	7.65e+00
c	36	24	0.6	43637	14	CEB0399	Caenorhabditis elegans	2.56e+01
c	37	24	0.6	79355	19	AB012247	Arabidopsis thaliana g	2.56e+01
c	38	24	0.6	84510	19	AB006698	Arabidopsis thaliana g	2.56e+01
c	39	24	0.6	112392	21	AC004028	Human PAC clone DJ0800	2.56e+01
	40	24	0.6	128598	12	D90912	Synechocystis sp. PCC6	2.56e+01
	41	24	0.6	130380	13	HS352E11	Human DNA sequence ***	2.56e+01
	42	24	0.6	132645	13	AC002345	*** SEQUENCING IN PROG	2.56e+01
c	43	25	0.6	139698	13	HS929C8	Human DNA sequence ***	7.65e+00
	44	25	0.6	156399	21	AC002456	Human BAC clone RG013L	7.65e+00
c	45	24	0.6	166983	21	AC002541	Human BAC clone RG043K	2.56e+01

# ALIGNMENTS

LOCUS ATU87794 5655 bp DNA PLN 25-MAY-1997  
 DEFINITION Arabidopsis thaliana transcription factor inhibitor I kappa B  
 homolog (nim1) gene, complete cds.  
 ACCESSION U87794  
 NID g1916911  
 KEYWORDS .  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryotae; mitochondrial eukaryotes; Viridiplantae;  
 Charophyta/Embryophyta group; Embryophyta; vascular plants; seed  
 plants; Magnoliophyta; Magnoliopsida; Capparales; Brassicaceae;  
 Arabidopsis.  
 REFERENCE 1 (bases 1 to 5655)  
 AUTHORS Ryals,J., Weymann,K., Lawton,K., Friedrich,L., Ellis,D.,  
 Steiner,H.-Y., Johnson,J., Delaney,T.P., Jesse,T., Vos,P. and  
 Uknes,S.  
 TITLE The Arabidopsis NIM1 protein shows homology to the mammalian  
 transcription factor inhibitor I kappa B  
 JOURNAL Plant Cell 9 (3), 425-439 (1997)  
 MEDLINE 97246324  
 REFERENCE 2 (bases 1 to 5655)  
 AUTHORS Ryals,J., Weymann,K., Lawton,K., Friedrich,L., Ellis,D.,  
 Steiner,H.-Y., Johnson,J., Delaney,T.P., Jesse,T., Vos,P. and  
 Uknes,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-JAN-1997) Fungicides, Novartis BGC, 3054 Cornwallis  
 Road, Research Triangle Park, NC 27709, USA  
 FEATURES Location/Qualifiers  
 source 1. .5655  
 /organism="Arabidopsis thaliana"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /map="between ATHGENEA and nga111"  
 mRNA join(2754. .3347,3427. .4162,4271. .4474,4586. .5053)  
 gene 2787. .4866  
 /gene="nim1"  
 CDS join(2787. .3347,3427. .4162,4271. .4474,4586. .4866)  
 /gene="nim1"  
 /codon\_start=1  
 /product="transcription factor inhibitor I kappa B  
 homolog"  
 /db\_xref="PID:g1916912"  
 /translation="MDTTIDGFADSYEISSTSFVATDNTDSSIVYLAAEQVLTGPDVS  
 ALQLLSNSFESVFDSPDDFYSDAKLVLSDBGREVSFHRVLSARSSFFKSALAAKKEK  
 DSNNTAAVKLELKEIAKDYEVGFDVTVLAYVYSSRVPPPKGVSECADENCCHVAC  
 RPAVDFMLEVLYLAFIFKIPELITLYQRHLLDVVDKVVIEDTLVILKLANICGKACMK  
 LLDRCKEIIVKSNDVMSLEKSLPEELVKEIIDRRKELGLEVPKVKKHVSNVHKALDS  
 DDIELVKLLKEDHTNLDACALHFAVAYCNVKTATDLLKLDLADVNHRNPRGYTVLH  
 VAAMRKEPQLILSLLEKGASASEATLEGR TALMIAKQATMAVECNNIPEQCKHSLKGR  
 LCVEILEQEDKREQIPRDVPPSFAVADELKMTLLDLENRVALAQRLFPTEAQAAMEI  
 AEMKGTCEFIVTSLEPDRLTGKRTSPGVKIAPFRILEEHQSRLKALSKTVELGKRFF  
 PRCSAVLDQIMNCEDLTQLACGEDDTAEKRLQKKQRYMEIQETLKKAFSEDNLELGNS  
 SLTDSTSSTSKSTGGKRSNRKLSHRRR"

BASE COUNT	1758 a	1014 c	1069 g	1814 t
ORIGIN				

Query Match 50.6%; Score 1949; DB 19; Length 5655;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1961; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db	3694	GATGATATTGAGTTAGTCAAGTTGCTTTTGAAAGAGGATCACACCAATCTAGATGATGCG	3753
Qy	3700	GATGATATTGAGTTAGTCAAGTTGCTTTTGAAAGAGGATCACACCAATCTAGATGATGCG	3759
Db	3754	TGTGCTCTTCATTTTCGCTGTTGCATATTGCAATGTGAAGACCGCAACAGATCTTTTAAAA	3813
Qy	3760	TGTGCTCTTCATTTTCGCTGTTGCATATTGCAATGTGAAGACCGCAACAGATCTTTTAAAA	3819
Db	3814	CTTGATCTTGCCGATGTCAACCATAGGAATCCGAGGGGATATACGGTGCTTCATGTTGCT	3873
Qy	3820	CTTGATCTTGCCGATGTCAACCATAGGAATCCGAGGGGATATACGGTGCTTCATGTTGCT	3879
Db	3874	GCGATGCGGAAGGAGCCACAATTGATACTATCTCTATTGGAAAAAGGTGCAAGTGCATCA	3933
Qy	3880	GCGATGCGGAAGGAGCCACAATTGATACTATCTCTATTGGAAAAAGGTGCAAGTGCATCA	3939
Db	3934	GAAGCAACTTTGGAAGGTAGAACCGCACTCATGATCGCAAAACAAGCCACTATGGCGGTT	3993
Qy	3940	GAAGCAACTTTGGAAGGTAGAACCGCACTCATGATCGCAAAACAAGCCACTATGGCGGTT	3999
Db	3994	GAATGTAATAATATCCCGGAGCAATGCAAGCATTCTCTCAAAGGCCGACTATGTGTAGAA	4053
Qy	4000	GAATGTAATAATATCCCGGAGCAATGCAAGCATTCTCTCAAAGGCCGACTATGTGTAGAA	4059
Db	4054	ATACTAGAGCAAGAAGACAAACGAGAACAAATTCCTAGAGATGTTCCCTCCTCTTTTGCA	4113
Qy	4060	ATACTAGAGCAAGAAGACAAACGAGAACAAATTCCTAGAGATGTTCCCTCCTCCTTTTGCA	4119
Db	4114	GTGGCGGCCGATGAATTGAAGATGACGCTGCTCGATCTTGAAAATAGAGGTATCTATCAA	4173
Qy	4120	GTGGCGGCCGATGAATTGAAGATGACGCTGCTCGATCTTGAAAATAGAGGTATCTATCAA	4179
Db	4174	GTCTTATTTCTTATATGTTTGAATTAAATTTATGTCTCTCTATTAGGAAACTGAGTGAA	4233
Qy	4180	GTCTTATTTCTTATATGTTTGAATTAAATTTATGTCTCTCTATTAGGAAACTGAGTGAA	4239
Db	4234	CTAATGATAACTATTCTTTGTGTCGTCCACTGTTTAGTTGCACTTGCTCAACGTCTTTTT	4293
Qy	4240	CTAATGATAACTATTCTTTGTGTCGTCCACTGTTTAGTTGCACTTGCTCAACGTCTTTTT	4299
Db	4294	CCAACGGAAGCACAAAGCTGCAATGGAGATCGCCGAAATGAAGGGAACATGTGAGTTCATA	4353
Qy	4300	CCAACGGAAGCACAAAGCTGCAATGGAGATCGCCGAAATGAAGGGAACATGTGAGTTCATA	4359
Db	4354	GTGACTAGCCTCGAGCCTGACCGTCTCACTGGTACGAAGAGAACATCACCGGGTGTAAG	4413
Qy	4360	GTGACTAGCCTCGAGCCTGACCGTCTCACTGGTACGAAGAGAACATCACCGGGTGTAAG	4419

Db	4414	ATAGCACCTTTTCAGAATCCTAGAAGAGCATCAAAGTAGACTAAAAGCGCTTTCTAAAACC	4473
Qy	4420	ATAGCACCTTTTCAGAATCCTAGAAGAGCATCAAAGTAGACTAAAAGCGCTTTCTAAAACC	4479
Db	4474	GGTATGGATTCTCACCCACTTCATCGGACTCCTTATCACAAAAAACAAACTAAATGATC	4533
Qy	4480	GGTATGGATTCTCACCCACTTCATCGGACTCCTTATCACAAAAAACAAACTAAATGATC	4539
Db	4534	TTTAAACATGGTTTTGTTACTTGCTGTCTGACCTTGTTTTTTTATCATCAGTGGAATC	4593
Qy	4540	TTTAAACATGGTTTTGTTACTTGCTGTCTGACCTTGTTTTTTT-ATCATCAGTGGAATC	4598
Db	4594	GGGAAACGATTCTTCCCGCGCTGTTCCGGCAGTGCTCGACCAGATTATGAACTGTGAGGAC	4653
Qy	4599	GGGAAACGATTCTTCCCGCGCTGTTCCGGCAGTGCTCGACCAGATTATGAACTGTGAGGAC	4658
Db	4654	TTGACTCAACTGGCTTGCGGAGAAGACGACACTGCTGA-GAAACGACTACAAAAGAAGCA	4712
Qy	4659	TTGACTCAACTGGCTTGCGGAGAAGACGACACTGCTGAAGAAACGACTACAAAAGAAGCA	4718
Db	4713	AAGGTACATGGAAATACAAGAGACACTAAAGAAGGCCTTTAGTGAGGACAATTTGGAATT	4772
Qy	4719	AAGGTACATGGAAATACAAGAGACACTAAAGAAGGCCTTTAGTGAGGACAATTTGGAATT	4778
Db	4773	AGGAAATTCGTCCCTGACAGATTCGACTTCTTCCACATCGAAATCAACCGGTGGAAAGAG	4832
Qy	4779	AGGAAATTCGTCCCTGACAGATTCGACTTCTTCCACATCGAAATCAACCGGTGGAAAGAG	4838
Db	4833	GTCTAACCGTAAACTCTCTCATCGTCGTCGGTGAGACTCTTGCCTCTTAGTGTAATTTTT	4892
Qy	4839	GTCTAACCGTAAACTCTCTCATCGTCGTCGGTGAGACTCTTGCCTCTTAGTGTAATTTTT	4898
Db	4893	GCTGTACCATATAATTCTGTTTTTCATGATGACTGTAAGTGTGTTATGTCTATCGTTGGCGT	4952
Qy	4899	GCTGTACCATATAATTCTGTTTTTCATGATGACTGTAAGTGTGTTATGTCTATCGTTGGCGT	4958
Db	4953	CATATAGTTTCGCTCTTCGTTTTGCATCCTGTGTATTATTGCTGCAGGTGTGCTTCAAAC	5012
Qy	4959	CATATAGTTTCGCTCTTCGTTTTGCATCCTGTGTATTATTGCTGCAGGTGTGCTTCAAAC	5018
Db	5013	AAATGTTGTAACAATTTGAACCAATGGTATACAGATTTGTAATATATATTTATGTACATC	5072
Qy	5019	AAATGTTGTAACAATTTGAACCAATGGTATACAGATTTGTAATATATATTTATGTACATC	5078
Db	5073	AACAATAACCCATGATGGTGTTACAGAGTTGCTAGAATCAAAGTGTGAAATAATGTCAAA	5132
Qy	5079	AACAATAACCCATGATGGTGTTACAGAGTTGCTAGAATCAAAGTGTGAAATAATGTCAAA	5138
Db	5133	TTGTTTCATCTGTTGGATATTTTCCACCAAGAACCAAAGAATATTCAAGTTCCTGAACT	5192
Qy	5139	TTGTTTCATCTGTTGGATATTTTCCACCAAGAACCAAAGAATATTCAAGTTCCTGAACT	5198
Db	5193	TCTGGCAACATTTCATGTTATATGTATCTTCCTAATTCTTCCTTTAACCTTTTGTAACG	5252

Qy	5199	TCTGGCAACATTCATGTTATATGTATCTTCCTAATTCTTCCTTTAACCTTTTGTAACCTCG	5258
Db	5253	AATTACACAGCAAGTTAGTTTCAGGTCTAGAGATAAGAGAACACTGAGTGGGCGTGTAAG	5312
Qy	5259	AATTACACAGCAAGTTAGTTTCAGGTCTAGAGATAAGAGAACACTGAGTGGGCGTGTAAG	5318
Db	5313	GTGCATTCTCCTAGTCAGCTCCATTGCATCCAACATTTGTGAATGACACAAGTTAACAAT	5372
Qy	5319	GTGCATTCTCCTAGTCAGCTCCATTGCATCCAACATTTGTGAATGACACAAGTTAACAAT	5378
Db	5373	CCTTTGCACCATTTTCTGGGTGCATACATGGAAACTTCTTCGATTGAAACTTCCCACATGT	5432
Qy	5379	CCTTTGCACCATTTTCTGGGTGCATACATGGAAACTTCTTCGATTGAAACTTCCCACATGT	5438
Db	5433	GCAGGTGCGTTTCGCTGTCACTGATAGACCAAGAGACTGAAAGCTTTCACAAATTGCCCTC	5492
Qy	5439	GCAGGTGCGTTTCGCTGTCACTGATAGACCAAGAGACTGAAAGCTTTCACAAATTGCCCTC	5498
Db	5493	AAATCTTCTGTTTCTATCGTCATGACTCCATATCTCCGACCACTGGTCATGAGCCAGAGC	5552
Qy	5499	AAATCTTCTGTTTCTATCGTCATGACTCCATATCTCCGACCACTGGTCATGAGCCAGAGC	5558
Db	5553	CCACTGATTTTGAGGGAATTGGGCTAACCATTTCAGAGCTTCTGAGTCCTTCTTTTTTGAT	5612
Qy	5559	CCACTGATTTTGAGGGAATTGGGCTAACCATTTCAGAGCTTCTGAGTCCTTCTTTTTTGAT	5618
Db	5613	GTCCTTTATGTAGGAATCAAATTCCTTCTGACTTGTGGAT	5655
Qy	5619	GTCCTTTATGTAGGAATCAAATTCCTTCTGACTTGTGGAT	5661

\*\*\*\*\*

M P S R C H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Oct 6 07:04:53 1998; MasPar time 555.70 Seconds  
942.362 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-1  
Description: (3700-7548) from US08908884.seq (3 of 3)  
Perfect Score: 3849  
N.A. Sequence: 3700 GATGATATTGAGTTAGTCAA.....GGTGTGGCTAGTAGAAGCTT 7548  
Comp: CTACTATAACTCAATCAGTT.....CCACACCGATCATCTTCGAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 10.057; Variance 5.467; scale 1.840

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
c	1	2827	73.4	9919	40	V04631	Arabidopsis thaliana	0.00e+00
	2	1949	50.6	5655	40	V04632	Arabidopsis thaliana	0.00e+00
	3	49	1.3	204	1	N81164	Base substituted E.co	3.60e-15
	4	41	1.1	91	9	Q51746	Oligonucleotide probe	4.11e-10
c	5	42	1.1	91	9	Q51746	Oligonucleotide probe	9.88e-11
c	6	40	1.0	204	1	N81164	Base substituted E.co	1.69e-09
	7	33	0.9	114	12	Q70465	Generic DNA sequence	2.44e-05
c	8	34	0.9	114	12	Q70467	Generic DNA sequence	6.45e-06
c	9	29	0.8	67	24	T14322	Primer used in the la	4.20e-03
	10	32	0.8	114	12	Q70467	Generic DNA sequence	9.05e-05
	11	30	0.8	114	12	Q70468	Generic DNA sequence	1.19e-03
	12	30	0.8	114	12	Q70469	Generic DNA sequence	1.19e-03
	13	29	0.8	114	12	Q70466	Generic DNA sequence	4.20e-03
c	14	32	0.8	114	12	Q70465	Generic DNA sequence	9.05e-05
c	15	32	0.8	114	12	Q70468	Generic DNA sequence	9.05e-05
c	16	31	0.8	114	12	Q70470	Generic DNA sequence	3.31e-04
c	17	31	0.8	114	12	Q70469	Generic DNA sequence	3.31e-04
c	18	30	0.8	114	12	Q70466	Generic DNA sequence	1.19e-03
c	19	29	0.8	114	12	Q70473	Generic DNA sequence	4.20e-03
c	20	31	0.8	178	32	T76405	Human endothelin-1 an	3.31e-04
	21	28	0.7	39	7	Q51787	Mixed oligonucleotide	1.46e-02
	22	28	0.7	114	12	Q70472	Generic DNA sequence	1.46e-02
	23	28	0.7	114	12	Q70470	Generic DNA sequence	1.46e-02
	24	27	0.7	114	12	Q70473	Generic DNA sequence	4.93e-02
	25	26	0.7	114	12	Q70471	Generic DNA sequence	1.64e-01
c	26	27	0.7	114	12	Q70472	Generic DNA sequence	4.93e-02



c	27	26	0.7	168	32	T76270	Human MDNCF antisense	1.64e-01
	28	26	0.7	178	32	T76405	Human endothelin-1 an	1.64e-01
	29	28	0.7	250	32	T76438	Substance P antisense	1.46e-02
	30	27	0.7	501	3	N50028	Sequence encoding new	4.93e-02
	31	27	0.7	501	3	N50033	Sequence encoding new	4.93e-02
	32	27	0.7	501	3	N50027	Sequence encoding new	4.93e-02
	33	26	0.7	501	3	N50029	Sequence encoding new	1.64e-01
c	34	26	0.7	501	3	N50030	Sequence encoding new	1.64e-01
	35	26	0.7	984	17	Q94336	Degenerate Alteromona	1.64e-01
c	36	25	0.6	82	21	T13610	DC43 TSAR library gen	5.29e-01
	37	25	0.6	498	3	N50034	Sequence encoding new	5.29e-01
	38	25	0.6	501	3	N50024	Sequence encoding new	5.29e-01
	39	25	0.6	501	3	N50032	Sequence encoding new	5.29e-01
	40	25	0.6	501	3	N50031	Sequence encoding new	5.29e-01
c	41	25	0.6	501	3	N50026	Sequence encoding new	5.29e-01
c	42	25	0.6	501	3	N50025	Sequence encoding new	5.29e-01
c	43	25	0.6	501	3	N50032	Sequence encoding new	5.29e-01
c	44	25	0.6	501	3	N50023	Sequence encoding new	5.29e-01
c	45	25	0.6	501	3	N50029	Sequence encoding new	5.29e-01

#### ALIGNMENTS

RESULT 1  
 ID V04631 standard; DNA; 9919 BP.  
 AC V04631;  
 DT 17-JUN-1998 (first entry)  
 DE Arabidopsis thaliana BAC-04 containing the NIM1 gene.  
 KW NIM1; noninducible immunity; disease resistance; plants;  
 KW SAR gene expression; ss.  
 OS Arabidopsis thaliana.  
 PN WO9749822-A1.  
 PD 31-DEC-1997.  
 PF 10-MAR-1997; E01218.  
 PR 10-JAN-1997; US-035022.  
 PR 21-JUN-1996; US-020272.  
 PR 30-AUG-1996; US-024883.  
 PR 13-DEC-1996; US-033177.  
 PR 27-DEC-1996; US-773559.  
 PA (NOVS ) NOVARTIS AG.  
 PI Delaney TP, Ellis DM, Friedrich LB, Johnson JE,  
 PI Lawton KA, Ryals JA, Weymann K;  
 DR WPI; 98-077185/07.  
 PT NIM1 gene which allows activation in plant of systemic acquired  
 PT resistance - useful to confer broad spectrum disease resistance in  
 PT plants, specifically crop plants, e.g. rice, wheat, barley, rye and  
 PT corn  
 PS Claim 4; Fig 14; 153pp; English.  
 CC The sequence is that of the 9.9kb region of clone BAC-04 containing  
 CC the NIM1 (noninducible immunity) gene. It may be used to confer a  
 CC broad spectrum disease resistance in plants, specifically crop plants,  
 CC e.g. rice, wheat, barley, rye and corn. The NIM1 gene can be used to  
 CC confer universal disease susceptibility to plant cells, and their  
 CC progeny. It can also be used in a screening method for identifying

CC compounds capable of inducing broad spectrum disease resistance in  
CC plants, while the plant cells, and their progeny, can be used to  
CC isolate a gene fragment which allows expression of broad spectrum  
CC disease resistance in plants, or to incorporate the resistant trait  
CC into plant lines through breeding.

SQ Sequence 9919 BP; 3150 A; 1826 C; 1737 G; 3206 T;

Query Match 73.4%; Score 2827; DB 40; Length 9919;  
Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
Matches 2864; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

```
Db      1  tgatcatgaattgcgtgtaggggttggttttaagatagggatgagctgaagaaggcgg 60
      |||
Cp 6567  TGATCATGAATTGCGTGTAGGGTTGTGTTTTAAAGATAGGGATGAGCTGAAGAAGGCGGT 6508
      |||

Db     61  ggactggtgttccattagagggcagcaaaagtgtgtagtacaagagattgagaaggacga 120
      |||
Cp 6507  GGACTGGTGTTCCATTAGAGGGCAGCAAAAGTGTGTAGTACAAGAGATTGAGAAGGACGA 6448
      |||

Db    121  gtatacgtttaaatgcatcagatggaaatgcaattggcgcgctcgggcagattgaataga 180
      |||
Cp 6447  GTATACGTTTAAATGCATCAGATGGAAATGCAATTGGTCGCGTCGGGCAGATTGAATAGA 6388
      |||

Db    181  agaacatggacttgtaaagataactaagtgtagttgggtccacatacttggtgttctatta 240
      |||
Cp 6387  AGAACATGGACTTGTTAAGATAACTAAGTGTAGTTGGTCCACATACTTGTTGTTCTATTA 6328
      |||

Db    241  agccggaaaacttcaacttgtaatttgcagcagaagagattgagtgctctgatcagggtac 300
      |||
Cp 6327  AGCCGGAAAACCTTCAACTTGTAATTTGCAGCAGAAGAGATTGAGTGTCTGATCAGGGTAC 6268
      |||

Db    301  aaccactctaacagcagagttgaaaagtttggtgacatgcttaaaacttcaaagctgcg 360
      |||
Cp 6267  AACCCACTCTAACAGCAGAGTTGAAAAGTTTGGTGACATGCTTAAAACTTCAAAGCTGCG 6208
      |||

Db    361  ggcagcagaacaggaagtaatcaaagatcagagtttcagagtattgcctaaactaattgg 420
      |||
Cp 6207  GGCAGCAGAACAGGAAGTAATCAAAGATCAGAGTTTCAGAGTATTGCCTAAACTAATTGG 6148
      |||

Db    421  ctgcatttcactcatctaattgggctacttggtggactgcaatatgagcttttccctaattcc 480
      |||
Cp 6147  CTGCATTTCACTCATCTAATGGGCTACTTGTGGACTGCAATATGAGCTTTTCCCTAATCC 6088
      |||

Db    481  tgaatttgcattccttcggtggcgcggttttgggcggtttccacagtccattgaagggtttca 540
      |||
Cp 6087  TGAATTTGCATCCTTCGGTGGCGCGTTTTGGGCGTTTCCACAGTCCATTGAAGGGTTTCA 6028
      |||

Db    541  aactgtagacctctgatcatagtggattcaaaagacttgaacggcaagtaccc-tatga 599
      |||
Cp 6027  AACTGTAGACCTCTGATCATAGTGGATTCAAAGACTTGAACGGCAAGTACCCCTATGA 5968
      |||

Db    600  aattgatgattt-cctcaggactcgacgctgatgattgctttttcccgcttgcccttccg 658
      |||
Cp 5967  AATTGATGATTTTCCTCAGGACTCGACGCTGATGATTG-TTTTCCCGCTTGC-TTTCCG 5910
      |||
```

Db	659	cttaccaaagaagtgtccactgatagttggcgttgggtttctcactaatatcagagagaag	718
Cp	5909	CTTACCAAAGAAGTGTCCACTGATAGTTGGCGTTGGTTTCTCACTAATATCAGAGAGAAG	5850
Db	719	gtaacacaaaggaaagacgtttgccctcgctccagtcctcaccgacatagttgctggt	778
Cp	5849	GTAACACAAAGGAAAGACGTTTGCTTCGTCTCCAGTCCTCACCCGGACATAGTTGCTGTT	5790
Db	779	attaacgaacccggatcactgtggcaagaaccttgggtctatcacaggttctgtctggat	838
Cp	5789	ATTAACGAACCCGGATCACTGTGGCAAGAACCCTGGGTCTATCACAGGTTCTGTCTGGAT	5730
Db	839	tgtttttgcttacaattccatgatatttttgagactacaacctggtgagccttgtgaag	898
Cp	5729	TGTTTTTGCTTACAATTCCATGATATTTTGGAGACTACAACCTGGTGAGCCTTGTGAAG	5670
Db	899	caggctggatccacaagtcagaaggaagaatttgattcctacataaaggacatcaaaaag	958
Cp	5669	CAGGCTGGATCCACAAGTCAGAAGGAAGAATTTGATTCTACATAAAGGACATCAAAAAG	5610
Db	959	aaggactcagaagctcggaaatggttagcccaattccctcaaaatcagtgggctctggct	1018
Cp	5609	AAGGACTCAGAAGCTCGGAAATGGTTAGCCCAATTCCCTCAAAATCAGTGGGCTCTGGCT	5550
Db	1019	catgaccagtggtcggagatatggagtcatgacgatagaaacagaagatttgagggcaat	1078
Cp	5549	CATGACCAGTGGTCGGAGATATGGAGTCATGACGATAGAAACAGAAGATTTGAGGGCAAT	5490
Db	1079	ttgtgaaagctttcagtctcttgggtctatcagtgacagcgaacgcacctgcacatgtggg	1138
Cp	5489	TTGTGAAAGCTTTCAGTCTCTTGGTCTATCAGTGACAGCGAACGCACCTGCACATGTGGG	5430
Db	1139	aagtttcaatcgaagaagtttccatgtatgcaccagaaatgggtgcaaaggattgttaac	1198
Cp	5429	AAGTTTCAATCGAAGAAGTTTCCATGTATGCACCCAGAAATGGTGCAAAGGATTGTTAAC	5370
Db	1199	ttgtgtcattcacaaatgttggatgcaatggagctgactaggagaatgcaccttacacgc	1258
Cp	5369	TTGTGTCATTACAAATGTTGGATGCAATGGAGCTGACTAGGAGAATGCACCTTACACGC	5310
Db	1259	ccactcagtgttctcttatctctagacctgaaactaacttgctgtgtaattcgagttaca	1318
Cp	5309	CCACTCAGTGTCTCTTATCTCTAGACCTGAAACTAACTTGCTGTGTAATTCGAGTTACA	5250
Db	1319	aaaggttaaaggaagaattaggaagatacatataacatgaatggttgccagaagttcaggg	1378
Cp	5249	AAAGGTTAAAGGAAGAATTAGGAAGATACATATAACATGAATGTTGCCAGAAGTTCAGGG	5190
Db	1379	aacttgaatattcttttgggttcttgggtggaaaatatccaacagatgaacaatttgacatt	1438
Cp	5189	AACTTGAATATTCTTTTGGTTCTTGGTGGAAAATATCCAACAGATGAACAATTTGACATT	5130
Db	1439	atttcacacttttgattctagcaactctgtaacaccatcatggggttattggttgatgtacat	1498

Cp	5129	AT TTCACACTTTGATTCTAGCAACTCTGTAACACCATCATGGGTATTGTTGATGTACAT	5070
Db	1499	aaatatataattacaaatctgtataaccattgggttcaaattgttacaacatttgtttgaagc	1558
Cp	5069	AAATATATATTACAAATCTGTATACCATTGGTTCAAATTGTTACAACATTTGTTTGAAGC	5010
Db	1559	acacctgcagcaataatacacaggatgcaaaacgaagagcgaaactatatgacgccaacg	1618
Cp	5009	ACACCTGCAGCAATAATACACAGGATGCAAAACGAAGAGCGAAACTATATGACGCCAACG	4950
Db	1619	atagacataaacagttacagtcacatcatgaaaacagaattatatggtacagcaaaaattac	1678
Cp	4949	ATAGACATAAACAGTTACAGTCATCATGAAAACAGAATTATATGGTACAGCAAAAATTAC	4890
Db	1679	actaagaggcaagagtctcaccgacgacgatgagagagtttacggttagacctctttcca	1738
Cp	4889	ACTAAGAGGCAAGAGTCTCACCGACGACGATGAGAGAGTTTACGGTTAGACCTCTTTCCA	4830
Db	1739	ccggttgatttcgatgtggaagaagtcgaatctgtcagggacgaatttcctaattccaaa	1798
Cp	4829	CCGGTTGATTTCGATGTGGAAGAAGTCGAATCTGTCAGGGACGAATTCCTAATTCCAAA	4770
Db	1799	ttgtcctcactaaaggccttcttttagtgtctcttgtatttccatgtacctttgcttcttt	1858
Cp	4769	TTGTCCTCACTAAAGGCCTTCTTTAGTGTCTCTTGTATTTCATGTACCTTTGCTTCTTT	4710
Db	1859	tgtagtcgtttct-cagcagtgctcgtcttctccgcaagccagttgagtcagtcctcaca	1917
Cp	4709	TGTAGTCGTTTCTTCAGCAGTGTCTCTTCTCCGCAAGCCAGTTGAGTCAAGTCCTCACA	4650
Db	1918	gttcataatctggtcgagcactgccgaacagcgcggaagaatcgtttcccagagttccac	1977
Cp	4649	G TTCATAATCTGGTCGAGCACTGCCGAACAGCGCGGAAGAATCGTTTCCCAGATTCCAC	4590
Db	1978	tgatgataaaaaaaaaacaaggtcagacagcaagtaacaaaaccatgtttaagatcattta	2037
Cp	4589	TGATGATAAAAAAA-CAAGGTCAGACAGCAAGTAACAAAACCATGTTTAAAGATCATTTA	4531
Db	2038	gttttggtttttgtgataaggagtcgatgaagtgggtgagaatccataccggttttaga	2097
Cp	4530	GTTTTGTTTTTTGTGATAAGGAGTCCGATGAAGTGGGTGAGAATCCATACCGGTTTTAGA	4471
Db	2098	aagcgcttttagtctactttgatgctcttctaggattctgaaagtgctatctttacacc	2157
Cp	4470	AAGCGCTTTTAGTCTACTTTGATGCTCTTCTAGGATTCTGAAAGGTGCTATCTTTACACC	4411
Db	2158	cggtgatgttctcttcgtaccagtgagacggtcaggctcgaggctagtcactatgaactc	2217
Cp	4410	CGGTGATGTTCTCTTCGTACCAGTGAGACGGTCAGGCTCGAGGCTAGTCACTATGAACTC	4351
Db	2218	acatgttcccttcatttcggcgatctccattgcagcttgtgcttccggttgaaaaagacg	2277
Cp	4350	ACATGTTCCCTTCATTTCCGCGATCTCCATTGCAGCTTGTGCTTCCGTTGGAAAAAGACG	4291

```

Db    2278  ttgagcaagtgcaactaaacagtgaggacacacaaagaatagttatcattagttcactcag 2337
      |||
Cp    4290  TTGAGCAAGTGCAACTAAACAGTGGACGACACAAAGAATAGTTATCATTAGTTCACTCAG 4231

Db    2338  tttcctaatagagaggacataaaattttaattcaaacatataagaaataagacttgatagat 2397
      |||
Cp    4230  TTTCTAATAGAGAGGACATAAATTTAATTCAAACATATAAGAAATAAGACTTGATAGAT 4171

Db    2398  acctctattttcaagatcgagcagcgtcatcttcaattcatcgggccgccactgcaaaaga 2457
      |||
Cp    4170  ACCTCTATTTTCAAGATCGAGCAGCGTCATCTTCAATTCATCGGCCGCCACTGCAAAAGA 4111

Db    2458  gggaggaacatctctaggaatttgttctcgtttgtcttcttgccttagtatttctacaca 2517
      |||
Cp    4110  GGGAGGAACATCTCTAGGAATTTGTTCTCGTTTGTCTTCTTGCTCTAGTATTTCTACACA 4051

Db    2518  tagtcggccttttgagagaatgcttgcattgctccgggatattattacattcaaccgccat 2577
      |||
Cp    4050  TAGTCGGCCTTTGAGAGAATGCTTGCATTGCTCCGGGATATTATTACATTCAACCGCCAT 3991

Db    2578  agtggccttgttttgcgatcatgagtgcggttctaccttccaaagttgcttctgatgcact 2637
      |||
Cp    3990  AGTGGCTTGTTTTGCGATCATGAGTGCGGTTCTACCTTCCAAAGTTGCTTCTGATGCACT 3931

Db    2638  tgcacctttttccaatagagatagtatcaattgtggctccttccgcatcgcagcaacatg 2697
      |||
Cp    3930  TGCACCTTTTTCGAATAGAGATAGTATCAATTGTGGCTCCTTCCGCATCGCAGCAACATG 3871

Db    2698  aagcaccttatatcccctcggattcctatggttgacatcggcaagatcaagttttaaaag 2757
      |||
Cp    3870  AAGCACCGTATATCCCCTCGGATTCCTATGGTTGACATCGGCAAGATCAAGTTTTTAAAG 3811

Db    2758  atctgttgcggtccttcacattgcaatatgcaacagcgaaatgaagagcacacgcatcatc 2817
      |||
Cp    3810  ATCTGTTGCGGTCTTCACATTGCAATATGCAACAGCGAAATGAAGAGCACACGCATCATC 3751

Db    2818  tagattggtgtgatcctcctttcaaaagcaacttgactaactcaatatcatc 2868
      |||
Cp    3750  TAGATTGGTGTGATCCTCTTTCAAAAGCAACTTGACTAACTCAATATCATC 3700

```

RESULT 2

```

ID    V04632 standard; DNA; 5655 BP.
AC    V04632;
DT    17-JUN-1998 (first entry)
DE    Arabidopsis thaliana NIM1 gene.
KW    NIM1; noninducible immunity; disease resistance; plants;
KW    SAR gene expression; ss.
OS    Arabidopsis thaliana.
FH    Key          Location/Qualifiers
FT    CDS          2787..4866)
FT          /*tag= a
FT          /product= "NIM1 protein"
FT    exon        2787..3347

```

```

FT          /*tag= b
FT          /number= 1
FT  exon    3427..4162
FT          /*tag= c
FT          /number= 2
FT  exon    4271..4474
FT          /*tag= d
FT          /number= 3
FT  exon    4586..4863
FT          /*tag= e
FT          /number= 4
PN  WO9749822-A1.
PD  31-DEC-1997.
PF  10-MAR-1997; E01218.
PR  10-JAN-1997; US-035022.
PR  21-JUN-1996; US-020272.
PR  30-AUG-1996; US-024883.
PR  13-DEC-1996; US-033177.
PR  27-DEC-1996; US-773559.
PA  (NOVS ) NOVARTIS AG.
PI  Delaney TP, Ellis DM, Friedrich LB, Johnson JE,
PI  Lawton KA, Ryals JA, Weymann K;
DR  WPI; 98-077185/07.
DR  P-PSDB; W23963.
PT  NIM1 gene which allows activation in plant of systemic acquired
PT  resistance - useful to confer broad spectrum disease resistance in
PT  plants, specifically crop plants, e.g. rice, wheat, barley, rye and
PT  corn
PS  Claim 2; Fig 15; 153pp; English.
CC  The sequence is that of the NIM1 (noninducible immunity) gene. It
CC  may be used to confer a broad spectrum disease resistance in plants,
CC  specifically crop plants, e.g. rice, wheat, barley, rye and corn.
CC  The NIM1 gene can be used to confer universal disease susceptibility
CC  to plant cells, and their progeny. It can also be used in a screening
CC  method for identifying compounds capable of inducing broad spectrum
CC  disease resistance in plants, while the plant cells, and their
CC  progeny, can be used to isolate a gene fragment which allows
CC  expression of broad spectrum disease resistance in plants, or to
CC  incorporate the resistant trait into plant lines through breeding.
SQ  Sequence 5655 BP; 1758 A; 1014 C; 1069 G; 1814 T;

```

```

Query Match          50.6%; Score 1949; DB 40; Length 5655;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1961; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

```

Db  3694 gatgatattgagttagtcaagttgcttttgaaagaggatcacaccaatctagatgatgcg 3753
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3700 GATGATATTGAGTTAGTCAAGTTGCTTTTGAAAGAGGATCACACCAATCTAGATGATGCG 3759

Db  3754 tgtgctcttcatttcgctggttgcatttgcaatgtgaagaccgcaacagatcttttaaaa 3813
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3760 TGTGCTCTTCATTTGCTGTTGCATATTGCAATGTGAAGACCGCAACAGATCTTTTAAAA 3819

Db  3814 cttgatcttgccgatgtcaaccataggaatccgaggggatatacggtgcttcatgttgct 3873

```

Qy	3820	 CTTGATCTTGCCGATGTCAACCATAGGAATCCGAGGGGATATACGGTGCTTCATGTTGCT	3879
Db	3874	gcgatgcggaaggagccacaattgatactatctctattggaaaaaggtgcaagtgcac	3933
Qy	3880	 GCGATGCGGAAGGAGCCACAATTGATACTATCTCTATTGGAAAAAGGTGCAAGTGCATCA	3939
Db	3934	gaagcaacttttgaaggtagaaccgcactcatgatcgcaaaacaagccactatggcggtt	3993
Qy	3940	 GAAGCAACTTTTGGAAGGTAGAACCGCACTCATGATCGCAAAACAAGCCACTATGGCGGTT	3999
Db	3994	gaatgtaataatatcccggagcaatgcaagcattctctcaaaggccgactatgtgtagaa	4053
Qy	4000	 GAATGTAATAATATCCCGGAGCAATGCAAGCATTCTCTCAAAGGCCGACTATGTGTAGAA	4059
Db	4054	atactagagcaagaagacaaacgagaacaaattcctagagatgttcctccctcttttgca	4113
Qy	4060	 ATACTAGAGCAAGAAGACAAACGAGAACAAATTCTTAGAGATGTTCTCCTCCTCTTTTGCA	4119
Db	4114	gtggcgccgatgaattgaagatgacgctgctcgatcttgaaaatagaggtatctatcaa	4173
Qy	4120	 GTGGCGGCCGATGAATTGAAGATGACGCTGCTCGATCTTGAAAATAGAGGTATCTATCAA	4179
Db	4174	gtcttattttcttatatgtttgaattaaatttatgtcctctctattaggaaactgagtga	4233
Qy	4180	 GTCTTATTTCTTATATGTTTGAATTAAATTTATGTCCTCTCTATTAGGAAACTGAGTGAA	4239
Db	4234	ctaataataactattctttgtgtcgtccactgttttagttgcacttgctcaacgtcttttt	4293
Qy	4240	 CTAATGATAACTATTCTTTGTGTCTGTCCTGTTTAGTTGCACTTGCTCAACGTCTTTTT	4299
Db	4294	ccaacggaagcacaagctgcaatggagatcgccgaaatgaagggaacatgtgagttcata	4353
Qy	4300	 CCAACGGAAGCACAAGCTGCAATGGAGATCGCCGAAATGAAGGGAACATGTGAGTTCATA	4359
Db	4354	gtgactagcctcgagcctgaccgtctcactggtacgaagagaacatcaccgggtgtaaag	4413
Qy	4360	 GTGACTAGCCTCGAGCCTGACCGTCTCACTGGTACGAAGAGAACATCACCGGGTGTAAAG	4419
Db	4414	atagcacctttcagaatcctagaagagcatcaaagtagactaaaagcgcttttctaaaacc	4473
Qy	4420	 ATAGCACCTTTCAGAATCCTAGAAGAGCATCAAAGTAGACTAAAAGCGCTTCTCTAAAACC	4479
Db	4474	ggtatggattctcaccacttcatcgactccttatcacaaaaacaaaactaaatgac	4533
Qy	4480	 GGTATGGATTCTCACCCACTTCATCGGACTCCTTATCACAAAAACAAAACATAATGATC	4539
Db	4534	tttaaacatggttttgttacttgctgtctgaccttgttttttttatcatcagtgggaactc	4593
Qy	4540	 TTTAAACATGGTTTTGTTACTTGCTGTCTGACCTTGTTTTTTT-ATCATCAGTGGAACCTC	4598
Db	4594	gggaaacgattcttcccgcgctgttcggcagtgctcgaccagattatgaactgtgaggac	4653
Qy	4599	 GGGAAACGATTCTTCCCGCGCTGTTTCGGCAGTGCTCGACCAGATTATGAACTGTGAGGAC	4658

Db	4654	ttgactcaactggcttgcggagaagacgacactgctga-gaaacgactacaaaagaagca	4712
Qy	4659	TTGACTCAACTGGCTTGCGGAGAAGACGACACTGCTGAAGAAACGACTACAAAAGAAGCA	4718
Db	4713	aaggtacatggaaatacaagagacactaaagaaggccttttagtgaggacaatttggaatt	4772
Qy	4719	AAGGTACATGGAAATACAAGAGACACTAAAGAAGGCCTTTAGTGAGGACAATTTGGAATT	4778
Db	4773	aggaaattcgtccctgacagattcgacttcttccacatcgaaatcaaccgggtggaaagag	4832
Qy	4779	AGGAAATTCGTCCCTGACAGATTCGACTTCTTCCACATCGAAATCAACCGGTGGAAAGAG	4838
Db	4833	gtctaaccgtaaactctctcatcgtcgtcggtgagactcttgccctcttagtgtaatTTTT	4892
Qy	4839	GTCTAACCGTAAACTCTCTCATCGTCGTCGGTGAGACTCTTGCCCTCTTAGTGTAATTTTT	4898
Db	4893	gctgtaccatataattctgttttcatgatgactgtaactgtttatgtctatcgttggcgt	4952
Qy	4899	GCTGTACCATATAATTCTGTTTTCATGATGACTGTAACTGTTTATGTCTATCGTTGGCGT	4958
Db	4953	catatagtttcgctcttcgttttgcatcctgtgtattattgctgcagggtgtgcttcaaac	5012
Qy	4959	CATATAGTTTCGCTCTTCGTTTTGCATCCTGTGTATTATTGCTGCAGGTGTGCTTCAAAC	5018
Db	5013	aatggttgtaacaatttgaaccaatggtatacagatttgtaatatatatttatgtacatc	5072
Qy	5019	AAATGTTGTAACAATTTGAACCAATGGTATACAGATTGTAAATATATATTTATGTACATC	5078
Db	5073	aacaataacccatgatggtgttacagagttgctagaatcaaagtgtgaaataatgtcaaa	5132
Qy	5079	AACAATAACCCATGATGGTGTACAGAGTTGCTAGAATCAAAGTGTGAAATAATGTCAAA	5138
Db	5133	ttgttcacatctgttggatattttccaccaagaaccaaagaatattcaagttccctgaact	5192
Qy	5139	TTGTTACATCTGTTGGATATTTTCACCAAGAACC AAAAGAATATTCAAGTTCCTGAACT	5198
Db	5193	tctggcaacattcatgttatatgtatcttcctaattcttcctttaaccttttgtaactcg	5252
Qy	5199	TCTGGCAACATTCATGTTATATGTATCTTCCTAATTCTTCCTTTAACCTTTTGTAACCTCG	5258
Db	5253	aattacacagcaagttagtttcagggtctagagataagagaacactgagtgggcgtgtaag	5312
Qy	5259	AATTACACAGCAAGTTAGTTTCAGGTCTAGAGATAAGAGAACACTGAGTGGGCGTGTAAG	5318
Db	5313	gtgcattctcctagtcagctccattgcatccaacatttgtgaatgacacaagttaacaat	5372
Qy	5319	GTGCATTCTCCTAGTCAGCTCCATTGCATCCAACATTTGTGAATGACACAAGTTAACAAT	5378
Db	5373	cctttgcaccatttctgggtgcatacatggaaacttcttcgattgaaacttcccacatgt	5432
Qy	5379	CCTTTGCACCATTTCTGGGTGCATACATGGAAACTTCTTCGATTGAAACTTCCCACATGT	5438
Db	5433	gcagggtgcgttcgctgtcactgatagaccaagagactgaaagctttcacaaattgccctc	5492



Qy 5439 GCAGGTGCGTTCGCTGTCACTGATAGACCAAGAGACTGAAAGCTTTCACAAATTGCCCTC 5498

Db 5493 aaatcttctgtttctatcgtcatgactccatatctccgaccactggatcatgagccagagc 5552  
|||||

Qy 5499 AAATCTTCTGTTTCTATCGTCATGACTCCATATCTCCGACCACTGGTCATGAGCCAGAGC 5558

Db 5553 ccactgattttgaggggaattgggctaaccatttccgagcttctgagtccttctttttgat 5612  
|||||

Qy 5559 CCACTGATTTTGAGGGAATTGGGCTAACCATTTCCGAGCTTCTGAGTCCTTCTTTTGAT 5618

Db 5613 gtcctttatgtaggaatcaaattcttccttctgacttgtggat 5655  
|||||

Qy 5619 GTCCTTTATGTAGGAATCAAATTCTTCCTTCTGACTTGTGGAT 5661

\*\*\*\*\*

MPsrch (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Oct 6 05:58:05 1998; MasPar time 3957.16 Seconds  
1362.257 Million cell updates/sec

Tabular output not generated.

Title: >US-08-908-884-1  
Description: (3700-7548) from US08908884.seq (3 of 3)  
Perfect Score: 3849  
N.A. Sequence: 3700 GATGATATTGAGTTAGTCAA.....GGTGTGGCTAGTAGAAGCTT 7548  
Comp: CTACTATAACTCAATCAGTT.....CCACACCGATCATCTTCGAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1832099 seqs, 700269816 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embl-est55

1:em\_est10 2:em\_est11  
Database: genbank-est107  
3:gb\_est1 4:gb\_est10 5:gb\_est11 6:gb\_est12 7:gb\_est13  
8:gb\_est14 9:gb\_est15 10:gb\_est16 11:gb\_est17  
12:gb\_est18 13:gb\_est19 14:gb\_est2 15:gb\_est20  
16:gb\_est21 17:gb\_est22 18:gb\_est23 19:gb\_est3  
20:gb\_est4 21:gb\_est5 22:gb\_est6 23:gb\_est7 24:gb\_est8  
25:gb\_est9 26:gb\_gss

Statistics: Mean 12.357; Variance 2.094; scale 5.901

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	377	9.8	470	11	AA395706	28108 Lambda-PRL2 Arab	0.00e+00
c	2	276	7.2	798	26	B21790	F7A9-Sp6 IGF Arabidops	0.00e+00
	3	235	6.1	258	11	T22612	4620 Lambda-PRL2 Arabi	0.00e+00
	4	208	5.4	909	26	B13248	F21L17-Sp6 IGF Arabido	0.00e+00
	5	154	4.0	783	26	B20373	F7D16-Sp6 IGF Arabidop	1.12e-279
	6	155	4.0	868	26	B20379	F7O9-Sp6 IGF Arabidops	4.86e-282
	7	154	4.0	936	26	B11275	F11H8-T7 IGF Arabidops	1.12e-279
	8	150	3.9	785	26	B20920	F22C12-T7 IGF Arabidop	3.01e-270
c	9	77	2.0	629	26	B26306	F18D8TF IGF Arabidopsi	6.04e-104
	10	71	1.8	252	13	AA754459	97SN1787 Rice Immature	4.28e-91
c	11	71	1.8	252	13	AA754459	97SN1787 Rice Immature	4.28e-91
	12	51	1.3	247	13	AA754458	97SN1784 Rice Immature	3.91e-50
c	13	47	1.2	247	13	AA754458	97SN1784 Rice Immature	2.01e-42
c	14	40	1.0	2275	12	AF034173	Homo sapiens ntcon2 co	1.71e-29
c	15	34	0.9	448	26	B24539	F20L24TF IGF Arabidops	3.71e-19
	16	35	0.9	2275	12	AF034173	Homo sapiens ntcon2 co	8.04e-21
c	17	26	0.7	448	14	R06759	yf11d04.s1 Homo sapien	5.93e-07
c	18	26	0.7	515	26	B28411	F18G19TF IGF Arabidops	5.93e-07
c	19	26	0.7	619	26	FR0004306	F.rubripes GSS sequenc	5.93e-07
	20	24	0.6	153	8	AA566154	MA3AS043.AAS S. manson	2.69e-04
c	21	23	0.6	214	6	HUM003B01A	Human fetal brain cDNA	4.81e-03
	22	24	0.6	309	21	ATTS1721	A. thaliana transcribe	2.69e-04
c	23	24	0.6	309	11	AA668490	ab86e12.s1 Stratagene	2.69e-04
	24	24	0.6	328	3	T85155	yd31d12.r1 Homo sapien	2.69e-04
c	25	24	0.6	353	13	AA799028	vn40c04.r1 Stratagene	2.69e-04
c	26	24	0.6	389	14	R06674	yf10g05.r1 Homo sapien	2.69e-04
	27	24	0.6	392	3	R00337	ye69d02.r1 Homo sapien	2.69e-04
c	28	24	0.6	393	15	AA856087	vw82g02.r1 Stratagene	2.69e-04
c	29	24	0.6	408	8	AA279720	zs92d01.r1 NCI_CGAP_GC	2.69e-04
c	30	25	0.6	424	13	AA742184	ny28c02.s1 NCI_CGAP_GC	1.33e-05
c	31	24	0.6	424	25	AA178808	ms15g08.r1 Stratagene	2.69e-04
	32	24	0.6	435	20	H66213	yu16h10.s1 Homo sapien	2.69e-04
	33	24	0.6	443	17	AA924097	UI-R-A1-du-h-06-0-UI.s	2.69e-04
c	34	23	0.6	452	7	AA447498	zw90h04.r1 Soares tota	4.81e-03

	35	25	0.6	457	20	H82453	yv80b01.r1 Homo sapien	1.33e-05
	36	24	0.6	481	17	C83144	Oryctolagus cuniculus	2.69e-04
c	37	24	0.6	481	17	C83011	Oryctolagus cuniculus	2.69e-04
	38	24	0.6	491	20	N28331	yx26a03.r1 Homo sapien	2.69e-04
	39	24	0.6	548	10	AA601299	no15d07.s1 NCI_CGAP_Ph	2.69e-04
c	40	25	0.6	571	9	AA557891	nl64g03.s1 NCI_CGAP_Pr	1.33e-05
c	41	23	0.6	582	26	B29097	T26E10TF TAMU Arabidop	4.81e-03
	42	24	0.6	660	12	AF034177	Homo sapiens ntcon6 co	2.69e-04
	43	23	0.6	1011	26	B09171	T2H15-Sp6 TAMU Arabido	4.81e-03
c	44	25	0.6	1119	26	B10269	F15H21-Sp6 IGF Arabido	1.33e-05
c	45	25	0.6	1287	13	AF038250	Homo sapiens clone ntc	1.33e-05

#### ALIGNMENTS

RESULT 1

LOCUS AA395706 470 bp mRNA EST 31-OCT-1997

DEFINITION 28108 Lambda-PRL2 Arabidopsis thaliana cDNA clone 9905XP 3', mRNA sequence.

ACCESSION AA395706

NID g2048931

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana  
Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;  
Embryophyta; vascular plants; seed plants; Magnoliophyta;  
Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 470)

AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.

TITLE Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones

JOURNAL Plant Physiol. 106, 1241-1255 (1994)

MEDLINE 95148729

COMMENT

Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, Mi  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@ibm.cl.msu.edu  
Seq primer: M13 -21 dye primer.

FEATURES

source Location/Qualifiers

1. .470  
/organism="Arabidopsis thaliana"  
/strain="var columbia"  
/note="Vector: lambda Zip-Lox; Site\_1: Sal; Site\_2: Not;  
Lambda PRL2 is a cDNA library derived from equal  
quantities of 4 pools of mRNA. The mRNA sources were 1) 7  
day germinated etiolated seedlings; 2) tissue culture  
grown roots; 3) staged plants half with 24 hour light

cycle, half on 16 hr light, 8 hour dark- rosettes; 4)  
 same plants as 3 but aerial tissue (stems, flowers and  
 siliques. The vector is BRL's lambda Zip-Lox. The cDNA  
 inserts were directionally cloned with Sal-Not arms using  
 oligo dT primed cDNA."

/db\_xref="taxon:3702"

/clone="9905XP"

/clone\_lib="Lambda-PRL2"

BASE COUNT        119 a        101 c        99 g        143 t        8 others  
 ORIGIN

Query Match                    9.8%;    Score 377;    DB 11;    Length 470;  
 Best Local Similarity    96.5%;    Pred. No. 0.00e+00;  
 Matches    409;    Conservative        0;    Mismatches    11;    Indels        4;    Gaps        4;

```

Db      49  CTCAACTGGCTTCCGGAGAGGNCGACCCTGCTGG-GAACCGNCTACCAAAGGAAGCAAAG 107
      |||
Qy     4663 CTCAACTGGCTTGCGGAGAAGACGACACTGCTGAAGAAACGACTACAAAAG-AAGCAAAG 4721

Db      108  GTCCATGGAA-TACAAGAGNCACTAAAGAAGGCCTTTAGTGAGGACAATTTGGAATTAGG 166
      |||
Qy     4722  GTACATGGAAATACAAGAGACACTAAAGAAGGCCTTTAGTGAGGACAATTTGGAATTAGG 4781

Db      167  AAATTCGTCCCTGACAGATTCGACTTCTTCCCACATCGAAATCAACCGGTGGAAAGAGGT 226
      |||
Qy     4782  AAATTCGTCCCTGACAGATTCGACTTCTTCC-ACATCGAAATCAACCGGTGGAAAGAGGT 4840

Db      227  CTAACCGTAAACTCTCTCATCGTCGTCGGTGAGACTCTTGCCTCTTAGTGTAATTTTGC 286
      |||
Qy     4841  CTAACCGTAAACTCTCTCATCGTCGTCGGTGAGACTCTTGCCTCTTAGTGTAATTTTGC 4900

Db      287  TGTACCATATAATTCTGTTTTCATGATGACTGTAAGTGTATGTCTATCGTTGGCGTCA 346
      |||
Qy     4901  TGTACCATATAATTCTGTTTTCATGATGACTGTAAGTGTATGTCTATCGTTGGCGTCA 4960

Db      347  CATAGTTTCGCTCTTCGTTTTGCATCCTGTGTATTATTGCTGCAGGTGTGCTTCAAACAA 406
      |||
Qy     4961  TATAGTTTCGCTCTTCGTTTTGCATCCTGTGTATTATTGCTGCAGGTGTGCTTCAAACAA 5020

Db      407  ATGTTGTAACAATTTGAACCAATGGTATACAGATTTGTAATATATATTTATGTACATCAA 466
      |||
Qy     5021  ATGTTGTAACAATTTGAACCAATGGTATACAGATTTGTAATATATATTTATGTACATCAA 5080

Db      467  CAAT 470
      |||
Qy     5081  CAAT 5084
  
```